

Reviews

The American Jury: A Reassessment

Michael H. Walsh†

The American Jury. By Harry Kalven, Jr. & Hans Zeisel. Boston & Toronto: Little, Brown & Co., 1966. Pp. xi, 559. \$15.00.

Debate over the merits of maintaining the American jury system has long been a favorite subject of legal controversy. The literature in the field is voluminous;¹ the discussion has been marked by bitterness and exaggeration.² On the basic issue of the jury's merit as a fact finder, both partisans and opponents argue that its distinctive characteristics evidence either the enduring strengths or the inherent weakness of the jury. Those who favor the system argue that twelve heads, even of moderate intelligence, are better than one even of superior intelligence. The collective memory and the common experiences of the jurors, the proponents suggest, help the jury to comprehend, assess, and recall what took place at trial. Those who oppose the system point

† B.A. 1964 Stanford University; LL.B. 1969 Yale University. This review was originally a Senior Studies paper submitted in the spring of 1969.

1. One commentator estimated three hundred articles have been written on the subject. Broeder, *Memorandum Regarding Jury System, Hearings on Recording of Jury Deliberations Before the Subcomm. to Investigate the Administration of the Internal Security Act of the Senate Comm. on the Judiciary*, 84th Cong., 1st Sess. 64 (1955).

2. See Seville, *Trial by Jury: An Ineffective Survival*, 10 A.B.A.J. 53, 55 (1924):

Too long has the effete and sterile jury system been permitted to tug at the throat of the Nation's judiciary as it sinks under the smothering deluge of the obloquy of those it was designed to serve. Too long has ignorance been permitted to sit ensconced in the places of judicial administration where knowledge is so sorely needed. Too long has the lament of the Shakespearean character been echoed, "Justice has fled to brutish beasts and men have lost their reason."

Or see LORD DEVLIN, *TRIAL BY JURY* 164 (1956):

Each jury is a little parliament. The jury sense is the parliamentary sense. I cannot see the one dying and the other surviving. The first object of any tyrant in Whitehall would be to make Parliament utterly subservient to his will; and the next to overthrow or diminish trial by jury, for no tyrant could afford to leave a subject's freedom in the hands of twelve of his countrymen. So that trial by jury is more than an instrument of justice and more than one wheel of the constitution: it is the lamp that shows that freedom lives.

Senator James O. Eastland, Chairman of the Subcommittee on Internal Security of the Committee on the Judiciary, defended the exercise of his subcommittee jurisdiction to investigate alleged jury "tapping" by members of the Chicago Jury Project: "The jurisdiction . . . arises from the fact that anything which undermines or threatens the integrity of the jury system necessarily affects the internal security of the United States." *Hearings*, *supra* note 1, at 1.

out that jurors are not selected in a manner designed to maximize their aggregate intelligence. Moreover, they operate in an emotion-charged atmosphere not especially conducive to truth-seeking and, in many jurisdictions, are not allowed to take notes on complicated subjects or directly to ask questions of the witnesses.

The problem, of course, is that both the above arguments are valid. Juries are not selected or allowed to function in a manner which maximizes their fact finding capacity. Yet, the assertion that twelve minds acting collectively are better than one acting alone is not inherently unreasonable. At this point, the issues are joined but not resolved. The dialogue concerning many other important aspects of the jury debate is equally inconclusive.³

An effective resolution of the issues framed in the traditional jury debate requires much more knowledge about how juries actually operate. We need to know whether juries find facts at least as well as judges, and, if not, whether the jury serves other offsetting values. For these reasons, the extensive empirical research conducted by the Chicago Jury Project promised a breakthrough in the jury debate. By applying the skills of social science to a major unresolved legal problem, the authors of this prolonged and massively financed study hoped to replace speculation and conjecture with analysis and fact. Whether or not this important purpose has been achieved is the subject of this review.

*The American Jury*⁴ is the Chicago Jury Project's foremost publication. When it first appeared, Professor Abraham Goldstein described it as "a graceful and sophisticated book . . . of unquestioned importance," a "detailed report on the extent to which judge and jury disagree with one another in criminal cases, the reasons for the disagreement, and the ways in which highly developed methodology may be brought to bear on that problem."⁵ Furthermore, "the details of the research design, and the statistical treatment of the data, are presented with remarkable lucidity. And the limitations of both data

3. For example, are jurors better able to judge the credibility of witnesses because they can better "relate" to a witness who, like the juror, is in strange and unfamiliar surroundings for the first time? Or is the judge, because of his ability and experience, in a better position to decide these matters? Are judges or juries more likely to be biased, corrupted or narrow-minded? Does jury "lawlessness" speed or impede legal reform? Do juries by their alleged capriciousness cause the law to fall into disrespect? Or, by keeping the administration of justice close to the people, do they form the very basis of that respect? Has the jury survived because it continues to play an important role—or is it an anachronistic relic which is waiting to be discarded?

4. H. KALVEN & H. ZEISEL, *THE AMERICAN JURY* (1966) (hereinafter cited as KALVIN & ZEISEL).

5. Goldstein, Book Review, 1 L. & Soc'y REV. 148, 149, 151 (1967).

and research design are, for the most part, set out with admirable candor."⁶ Judge Henry Friendly was even more enthusiastic. He wrote that *The American Jury* is "a pioneering work . . . the sort of book that appears once in a decade."⁷ It is

a book to be savored and reread, not one to be gulped at a single sitting. Brilliantly avoiding Professor T. R. Powell's barb at the kind of research where "counters don't think and thinkers don't count," it shows how, in the hands of imaginative scholars and skillful writers, figures can enrich old insights and afford new ones.⁸

Although this praise was punctuated occasionally by polite criticism about such matters as the limitations of the sample or the limited objectives of the study itself, the reviewers agreed that *The American Jury* represented an important contribution to realist jurisprudence.

A reassessment of *The American Jury* suggests that this conclusion is unjustified. Kalven and Zeisel's data reveal little more than that judge and jury reach a different verdict in one of every four cases. Serious flaws in methodology gravely weaken the study's major conclusions. As a result, the jury debate remains almost as nonempirical as it was prior to the publication of *The American Jury*.

I.

To answer the principal question, "when do trial by judge and trial by jury lead to divergent results," the authors studied reports of some 3,576 criminal trials held in the United States between 1954 and 1958. Questionnaires were returned by 555 Federal and state court judges.⁹ Each judge was asked to report the jury's actual verdict and, more significantly, to indicate the decision he would have reached had he tried the case alone. The judges were also asked to supply descriptive and evaluative material about the case, the parties, and counsel.¹⁰

The authors found that in 75.4% of the cases jury and judge reached the same verdict, acquitting in 13.4% of the cases and convic-

6. *Id.* 148.

7. Friendly, Book Review, 33 U. CHI. L. REV. 884 (1966).

8. *Id.* 884-85.

9. For a full methodological breakdown, see KALVEN & ZEISEL, ch. 3. Briefly, of the 3500 judges originally approached, 555 effectively cooperated. The remainder did not respond, had no criminal trials, or did not cooperate for one reason or another. Half of the cases were submitted by 15% of the judges.

10. *Id.* 45.

ting in 62.0%. Almost all of the other conclusions of the study are based on the 25% of the cases in which the judge and jury disagreed. In cases of disagreement, the jury was consistently more lenient than the judge.¹¹ The authors admit the figures are subject to conflicting interpretations. The basic issue—how much disagreement is too much—is one for which “we lack a pre-existing context in which to place the measurements.”¹² They observe:

[t]o some, no doubt, the fact that judge and jury agree some 75% of the time will be read as a reassuring sign of the competence and stability of the jury system; to others the fact that they disagree 25% of the time will be viewed as a disturbing sign of the anarchy and eccentricity of the jury.¹³

Other questions¹⁴ permitted Kalven and Zeisel to assign highly specific reasons to the individual instances of judge-jury disagreement.¹⁵ Analyzing these reasons, they first combined similar cases in narrow sub-categories, then expanded the sub-categories into larger categories. In the final stage of the “coding process,” the data was organized under five Reason Categories—Evidentiary Factors, Jury Sentiments About the Law, Jury Sentiments About the Defendant, Disparity of Counsel, and Facts Only the Judge Knew.¹⁶ On the basis of these Reason Categories, the authors concluded that “unless at least one of these factors is present in a case, the jury and judge will not disagree.”¹⁷ By weighting the frequency with which Reason Categories were mentioned, they calculated that the causes of disagreement occurred with the following frequency:¹⁸

Sentiments on the law	29%
Sentiments on the defendant	11%
Issues of evidence	54%
Facts only the judge knew	2%
Disparity of counsel	4%
	<hr/> 100%

As a final step, the authors arranged their data in another profile.

11. *Id.* 61 n.8.

12. *Id.* 55.

13. *Id.* 57.

14. Actually the authors used two different questionnaires. The first asked longer, more open-ended questions; the second shorter, more specific ones. Except when Sample II “re-cruited special information,” the authors do not distinguish between the results of the two surveys. *Id.* 47-49.

15. *Id.* 105.

16. *Id.* 106.

17. *Id.* 109.

18. *Id.* 115.

By placing the evidence category to one side and combining the remaining four Reason Categories on the other, they classified the cases by disagreement about facts and disagreement about values. The results are summarized in Table 30:

<i>Disagree On</i>	<i>Per Cent</i>	
Facts alone	34	} } 79%—Total facts 66%—Total values
Values and Facts	45	
Values alone	21	

Thus formulated, the data led the authors to the central conclusion of their study: in nearly half the cases in which disagreement is found, juries, under the guise of resolving difficult factual issues, are giving rein to sympathies which the judge does not share. If the authors are correct, we know "in the large, but with precision, the answer to the question: what causes jury and judge to disagree."¹⁹ From other answers on the questionnaire, the authors conclude that the jury's importation of values meets with the judges' approval.²⁰

II.

Unfortunately, the authors' data will not support the wide-ranging interpretation argued for in *The American Jury*. Serious skepticism is justified in light of the inadequacy of the sample and questionnaire, reliance on "reason assessment," and the authors' use of unpersuasive analogies and deductions.

For a variety of reasons, the sample used was not statistically ideal. There was a distinct geographical imbalance in the trials reported. No control was maintained over the time period in which the cases being compared were tried. Judges were asked to give hypothetical verdicts and thus, arguably, were free to be more "rule-minded" than if human freedom had really been at stake. Almost complete reliance was placed on self-selection of participating judges. "Extra help or contacts to secure participation from a particular state"²¹ were also employed. A small minority of the participating judges (15%), a group already but a fraction of the judges contacted, submitted one-half of the responses. For the most part, however, the authors address

19. *Id.* 115

20. *Id.* 432-33.

21. *Id.* 36.

themselves frankly to these relatively minor difficulties. While such defects are "theoretically impressive,"²² they alone do not account for *The American Jury's* basic failings.

The authors' questionnaire raises much more substantial problems. It was the sole instrument of communication between the 555 participating judges and the authors. On the adequacy of the questionnaire, the answers it elicited from participating judges, and the authors' interpretation thereof, rests the entire validity of the conclusions drawn in *The American Jury*.

Of paramount importance in this regard is the nature of the questions asked. The overwhelming majority of the questions were *descriptive*. For example, the judge was asked, "As a person did [the defendant] evoke sympathy?," or "Were there extenuating circumstances in this case? (no . . . somewhat . . . decidedly . . .)," or "From the factual evidence . . . was the defendant's guilt or innocence very clear? A close question . . . ?," or "What is the community's sense of justice with respect to this type of crime? (In complete accord with the letter of the law . . . too severe . . . not severe enough?)"²³ In only one question was the issue of what motivated the jury to disagree with the judge specifically addressed. Question 34 read: "If you disagree with the jury, what in your opinion was the *main reason for the jury's verdict?*"(original emphasis).

Kalven and Zeisel did not, however, rely solely on answers to Question 34 to explain judge-jury disagreements.²⁴ In cases of disagreement they also inferred reasons from the answers to essentially descriptive questions. Thus, if the question "as a person did [the defendant] evoke sympathy?" were answered in the affirmative, for example, the authors listed as one reason for the disagreement "Jury Sentiments about the Individual Defendant." Similarly, if the question "what is the community's sense of justice with respect to this type of crime?" were answered "regards the law as *too severe*," one vote would be cast for "Jury Sentiments on the Law." No tally was kept as to which reasons were a product of inference from the answers to descriptive question and which resulted from the judge's response to the one specific question which inquired as to the "main reason for the jury's verdict."

There is a crucial difference between the question "As a person

22. *Id.* ch. 3.

23. The questionnaire is reprinted in *id.* 527.

24. Question 34 asked for the main reason for disagreement; in over half of the cases, the authors found more than one reason category involved.

did [the defendant] evoke sympathy?" and the question, "In your opinion, did the defendant evoke such sympathy in the mind of the jury as to influence them away from the weight of the evidence." The former question, in addition to being somewhat ambiguous (in whose mind did the defendant evoke sympathy—judge or jury or both?) is simply descriptive. The latter is a specific, evaluative question sharply focused on the issue about which the authors wished the judge to respond. To ask the former question and then treat the response as if it had been to the latter question is seriously to mislead the reader as to the basis upon which the authors' conclusions are founded.

Also disturbing are the means used by the authors to determine if disagreement resulted from a contradictory resolution of a disputed factual issue. The authors asked in question 12:

From the factual evidence in the case was the defendant's guilt or innocence

1) very clear:

2) a close question whether or not he was guilty beyond a reasonable doubt?

Evidentiary disagreement in these terms refers simply to the closeness of the case. Moreover, since the judge could only choose between "very clear" and "close," the only common characteristic of all the cases included in the latter category is that they were other than very clear on the evidence.

Not surprisingly, by this measure, the authors found evidentiary disagreement in four out of five cases of disagreement.²⁵ The natural follow-up question remains unasked, e.g., "In your opinion, was an evidentiary disagreement *responsible* for the verdict?"

To use but one additional example, the inability of the questionnaire to detect jury sentiments on the law should be noted. The only question directly dealing with this issue—which subsequently became a Reason Category all by itself—is Question 15:

What is the community's sense of justice with respect to this type of crime? (emphasis supplied)

1) In complete accord with the letter of the law

2) Regards the law as *too severe*

3) Regards the law as *not severe enough*

4) Don't know

The authors equated community sentiments with those of the

25. *Id.* 111.

jury. Thus, the answer "regards the law as *too severe*" led the authors to the conclusion that the disagreement was caused by "Jury Sentiments About the Law." This conclusion can be challenged on two grounds. Although the jury is theoretically representative of the community, the community's sense of justice and that of an individual jury (or juror) is not necessarily the same. Similarly, it is not unreasonable to suggest that in most cases where the jury regards the law as too severe, they will nevertheless honor their sworn duty to uphold the law as given by the judge. Hence, their sentiment as to the severity of the law will not necessarily lead to disagreement in result. It is clear that the preferable question would have been: "Was the jury's sense of justice with respect to this type of crime so much at odds with the letter of the law as you instructed them, as to influence them away from the weight of the evidence?"

In the 10% of the questionnaires in which the judge did not answer either Question 34 or a sufficient number of descriptive questions to allow the authors to infer an explanation, the authors used the factual material supplied by the judge in his preliminary description of the case to provide reasons for disagreement. Kalven and Zeisel claim that the omission was often a matter of form or that the facts clearly pointed to an explanation, even if the judge himself was mystified. They conclude:

In the end there are two things to emphasize about our intrusion into the assessment process. First, the clues always came from something the judge himself had said and considered worth reporting. Second, in view of our training and intense experience with the questionnaires in the study, we acquired some degree of expertise for the very special problem at hand.²⁶

There is no evidence, however, to suggest that the judge considered helpful facts offered in briefing the case as relevant to the issue of what influenced the jury. Nor do the elementary facts in the judge's brief provide any insight into the unreported facts which persuaded the judge not to answer a particular question. These other factors should carry as much weight as inferences drawn from factual statements about the case written in a context unrelated to the issue of judge-jury disagreement.

It should be emphasized that it makes a crucial difference what kind of questions the authors were asking. Without any effort, moreover, to compare the judges' descriptions in agreement versus disagreement

26. *Id.* 97.

cases, Kalven and Zeisel never eliminated the possibility that the characteristics they assigned to disagreement cases appeared in roughly equal frequency in cases of agreement.

Despite the nature of the questionnaire the authors represent themselves as having asked the judge to supply specific reasons for disagreement. This is the basis of "reason assessment," the name which the authors gave to their methodology. As they define it, reason assessment is

the assessment by a third party on an individual case by case basis. It was the trial judge who was the third party assessor as he gave his reasons for the causes of disagreement with the jury in the case before him.²⁷

In the use of this technique lies the basic weakness of *The American Jury*. The authors anticipated attack on this fundamental aspect of their work, but their attempt to explain away their difficulties is unconvincing.

The authors admit that both the preferred technique and their original strategy looked to a major use of "cross tabulation." This simply means that important variables are isolated, *i.e.*, everything but the variable is held constant. This is done by grouping cases so that each group shares as many common factual traits as possible. Differences in outcome may then be attributed to traits not commonly shared. To employ cross tabulation a massive amount of information is required so that cases may be accurately grouped and the distinguishing characteristics identified. Unfortunately, the authors' questionnaire did not yield sufficient information to make this sort of study possible. As the authors state:

The most surprising fact about the methodology of this study is that our apparently secure expectations of relying on cross-tabulation were largely, although not entirely, defeated.²⁸

Thus the authors decided that "a different method was required to locate the needed explanations." They invented "reason assessment," a "custom-made method with little precedent." General descriptive questions suddenly became highly specific, evaluative questions.

To defend "reason assessment," Kalven and Zeisel argue that there at least two methods that are analogous: the assessment of

27. *Id.* 92.

28. *Id.* 91.

causes in studies of traffic accidents and motivational research interviewing

. . . . [T]he novelty of our approach is that we combine the two methods.²⁹

The first analogy is ill-conceived. As the authors explain, in post-accident investigations the expert's familiarity with "the performance of automobiles, roads, and drivers, and the special configuration of post-accident clues, yields distinctive insights into the particular case."³⁰ Unlike the accident expert who deals with a limited number of physical phenomena which are preserved for his careful and extended scrutiny, a responding judge needed to cope with complex factual and personal variables³¹—all within a few moments at the end of a long trial. The authors acknowledge, moreover, that they do not know how many judges, if any, actually interviewed jurors before answering the questionnaire.³² Furthermore, if the judges followed the authors' instructions and turned their attention to the questionnaire before the jury returned, they lacked the most important datum of all—the jury verdict.

The second analogy is equally inapposite. Motivational research "seeks to establish reasons for individual acts by eliciting and analyzing the *actor's* motives as spelled out in his own utterances."³³ For relatively simple decisions it has proven to be a "powerful research tool;" but, as the authors state, "when the decision is a complex one, the method quickly reaches its limits."³⁴ As the authors learned in attempting to employ cross-tabulation, the jury process is indeed complex. More serious is the fact that there is no evidence that the authors had even indirect access to the jury's own statements. Further, the authors admit that motivational research depends on the actor's knowledge of his own motivations and on his willingness to disclose them.³⁵ Yet, as the authors acknowledge,³⁶ it is doubtful whether jurors fully understand their own motivations, and it is clear, in any event, that they did not disclose them.

The authors do not, however, rest their case on two misplaced analogies. There are other "strong reasons why the trial judge can be

29. *Id.* 92-93.

30. *Id.* 92.

31. *See id.* 91 for an explanation of why cross-tabulation did not work.

32. *Id.* 47-48 n.6.

33. *Id.* 92.

34. *Id.* 93.

35. *Id.* 93.

36. *Id.* ch. 38.

relied upon to locate the sources of judge-jury disagreement."³⁷ Having the benefit of "continuous exposure to jury behavior" in general and to the individual case, the judge is, Kalven and Zeisel argue, a "very special observer,"³⁸ who speaks with special authority about why the jury disagreed with him.

The problem with this argument is that judges' opinions were not subject to any corroboration. To illustrate the element of speculation inherent in the authors' treatment of the judges' replies, one might consider whether another judge and experienced counsel in the cases studied by the authors could agree on what influenced the jury to disagree with the judge, or whether two judges, independently observing the same case, could agree on what influenced the jury in reaching its ultimate conclusions. Furthermore, the expertise possessed by experienced counsel and the bench may lead to a prediction of results, but not to an explanation of the causes of disagreement.

Kalven and Zeisel also contend that since only cases of disagreement were considered, the inquiry "gains a helpful directionality [O]ne may expect in these cases a higher visibility of the jury's motives."³⁹ The obvious question is, Why? Jury deliberations were equally secret in cases of agreement and disagreement. Scrutinizing jurors in search of a clue to the inner workings of their minds amounts at best to a guess whose probability of success would not seem markedly changed by the fact that ultimately the judge and jury reach different conclusions. Furthermore, the judge did not know that the authors would only consider cases of disagreement. Presumably the judge tried equally hard to answer the authors' questions in all cases. Why he should be more believable in one case than another is not immediately apparent.

The authors' last argument is that the reasons themselves are the best evidence of the judge's ability to explain disagreements. Some of the explanations are "totally obvious on the face of the questionnaire." The judge often "has little difficulty empathizing with the jury." His reasons are rarely "given in terms of general stereotypes;" and the study reveals "highly individualized explanations for particular cases [which] tend to fall into larger categories of explanation."⁴⁰

The fact that judges were specific in their description is, of course, no proof that they were accurate. Expressed differently, an explanation's obviousness on the face of the questionnaire is no evidence that

37. *Id.* 94.

38. *Id.* 95.

39. *Id.* 95.

40. *Id.* 95-96.

it is correct. The assertion that the judge empathized with the jury in an unspecified number of disagreements only suggests that the judge may have believed what he was saying. That the explanations tended to fall into larger categories is simply evidence of the specificity of the questionnaire and the breadth of the authors' categories; it is certainly not evidence of the correctness of the judges' perceptions.

To gain a further insight into the accuracy of the judge's assessment of what influenced the jury, the reviewer conducted an admittedly artificial and limited empirical study of his own. Following each of seven mock Yale Law School criminal trials, the judge and each of the jurors was given a modified *American Jury* questionnaire. The jury completed the questionnaires prior to their deliberations (which were tape recorded). Thus the perceptions of the judge *and* the jury as to the characteristics of the case, of the parties, and of counsel could be compared. An analysis of the questionnaires reveals that judge and jury disagreed on half the questions. The jury generally thought the cases were harder and closer than did the judge. They were more likely to perceive extenuating circumstances. Contrary to *The American Jury's* findings, the jury was much less likely than the judge to perceive the defendant as "evoking sympathy." A wide disparity appeared in the assessment of the effectiveness of counsel. The judges found an imbalance in only one case, the jury in all but one. The percentage and direction of disagreement on guilt or innocence was almost identical to *The American Jury*. Judge and jury agreed in five of seven cases, and in both cases of disagreement, the jury was more lenient than the judge. Nonetheless, the serious deviation in the judge's and jury's perception of the case is further evidence that the basic assumption of *The American Jury* is seriously suspect. Rather than explaining disagreements in terms of sympathy for the defendant or against the law, the questionnaires and the tape recorded deliberations reveal that the juries were primarily concerned with the matter of reasonable doubt.

III.

In addition to presenting an explanation, however questionable, for disagreement between judges and juries, Kalven and Zeisel claim that their data show that juries are capable fact finders in that they render verdicts which follow the weight and direction of the evidence. The authors acknowledge that their study is not designed to test the fact-finding ability of the jury, but characterize their conclusions as a beneficial by-product of their broad research design.⁴¹

41. *Id.* 151-52.

The authors advance four arguments to support their contention that "the jury does by and large understand the facts and get the case straight." First they contend that agreement is too frequent to be consistent with the "hypothesis of substantial jury misunderstanding."⁴² If the jury does not understand, agreement with the judge would be a matter of chance; that is, it will occur in half the cases. Agreement in three out of four cases studied makes it improbable that much agreement was produced by chance. Furthermore the disagreement is highly directional, "thus compelling the conclusion that misunderstanding cannot in and of itself be a major factor in causing judge-jury disagreement."⁴³

That disagreement is directional hardly compels the conclusion that the jury understands—it may suggest the opposite. Confusion and misunderstanding often generate doubts which, given the presumption of innocence, tend to benefit the defense. This is, of course, the "direction" of the disagreement found by the authors.

That judge and jury agree in three-quarters of the cases also does not prove the jury understands. This conclusion presumes that the judge is always right. Suppose, moreover, that the jury understands half the time and misunderstands the other half. Assuming that in the half which they understand, they agree with the judge and that in the other half agreement is a product of chance, a total agreement of 75% would result, which is precisely what the study reveals.

Second, in 90% of the disagreement cases, the authors found a reason which constituted a "plausible explanation" for the disagreement. This, the authors suggest, "by its nature, precludes the notion that the jury did not understand the case."⁴⁴ Furthermore, with the exception of white-collar crime, *e.g.*, embezzlement, the judge "almost never advances the inability of the jury to understand as a reason for disagreement."⁴⁵

The fact that the judge seldom proffered the jury's inability to understand as a reason for disagreement is more a comment on the authors' questionnaire than on the content of the judge's responses. (The judge was never specifically asked if the jury understood the case, or if he believed their failure to do so was a reason for judge-jury disagreement.)

42. *Id.* 152.

43. *Id.* 152.

44. *Id.* 152.

45. *Id.* 152-53.

Third, the authors suggest that the large majority of criminal cases are easy to comprehend. The judge was asked the following question:

Compared to the average criminal case, was the evidence as a whole—
easy to comprehend?
somewhat difficult?
very difficult to comprehend?

Responses indicated that judges rated 86% “easy,” 12% “somewhat difficult,” and only 2% as “very difficult.” Thus, the authors conclude, “the great bulk of cases are routine as to comprehension and hence unlikely to be misunderstood.”⁴⁶

That 86% of the cases studied were rated by the judge as being of the same level of difficulty says absolutely nothing about how the average jury views the evidence in the day-to-day criminal case. For all that is known (referred to, *supra*, and the limited empirical study this suggests), the jury may find the average criminal case difficult to follow.

Fourth, Kalven and Zeisel argue for jury comprehension from the fact that the jury does not disagree with the judge any more frequently in cases it perceives as difficult than in cases it perceives as easy. If the jury has a tendency to misunderstand, it will most likely do so, and thus agree with the judge only by chance in the cases it perceives as difficult. The disagreement rate, however, turns out to be about the same in the “hard” as in the “easy” cases. This, according to Kalven and Zeisel, is “a stunning refutation of the hypothesis that the jury does not understand.”⁴⁷

The authors’ conclusions are questionable since the indicators of jury perception they employ—the frequency with which the jury returns to the judge for guidance, and the length of jury deliberations—are unreliable. As regards the jury return rate, the authors note that it is “surprising how infrequently” the jury returns to the judge for help. Furthermore, most inquiries are of a *legal* nature.⁴⁸ Thus the jury “return rate” is a more accurate indicator of the jury’s difficulty with the law rather than with the evidence in a given case. Moreover, the authors’ examples “of (evidentiary) questions that bother the jury” do not lead one to conclude that such questions accurately indicate the

46. *Id.* 154.

47. *Id.* 157.

48. *See id.* 511, Table 144.

difficulty of the case.⁴⁹ The argument derived from jury deliberation time is more substantial, but whether this one statistic supports the weight of the authors' conclusion is dubious.

The proposition that the jury is a capable fact-finder is indirectly supported by the finding that verdicts move with the weight and direction of the evidence.⁵⁰ The authors' figures may suggest movement in the right general direction but, assuming *arguendo* that the judge is always right, they are also consistent with a considerable amount of jury confusion.

The evidence produced by *The American Jury* on the jury's reliability as fact-finder, is, therefore, only tentative. The study was not designed to illuminate the problem, and the authors' assertions are based upon often unjustified inferences from the data.

IV.

Kalven and Zeisel assert that the responses to their questionnaire constitute "the judge's decisive ballot on how well the jury system is performing."⁵¹ As the authors interpret the data, the striking point is that the judge is critical of the jury's performance in only 9% of the total cases and only one-third of the disagreement cases. The judge is critical of only 15% of the Fact disagreements, 25% of the Fact-Value disagreements, and 78% of the Values alone disagreements. From these figures, the authors conclude that "for the judge as well as for the jury, evidentiary ambiguity legitimates the importation of values."⁵²

49. See *id.* 511; the examples are requests for more testimony, a dictionary, knowledge of the defendant's record, and the reason a witness had failed to give an item of information.

50. The authors' conclusion that the jury acquits more frequently in cases favorable to the defendant and convicts more frequently in cases favorable to the prosecution is based upon a combination of the judge's hypothetical verdict and his observation of whether the case was "clear" or "close." *Id.* 159, Table 51.

JURY ACQUITTAL IN CLEAR AND CLOSE CASES

	1 Clear cases- judge acquits	2 Close cases- judge acquits	3 Close cases- judge convicts	4 Clear cases- judge convicts
Jury acquits	95%	74%	46%	10%
Number of Cases	60	142	374	615

It is fair to ask which is the more impressive figure—that the jury conviction rate rises from 26% to 54% from category 2 to category 3—or that the disagreement rate in the two categories combined is 34%.

51. *Id.* 431.

52. *Id.* 433.

But do the figures tell us what Kalven and Zeisel claim?
In Sample II the judge was asked the following question:

5. Do you feel that the jury's verdict was (check one)
- 1) without any merit?
 - 2) a tenable position for a jury to take, though *not for a judge*?
 - 3) one a *judge* might also come to? (original emphasis)
 - 4) quite correct?

The results were:

Jury verdict was—	
Without merit	9%
Tenable for a jury, untenable for a judge	14%
One a judge might come to	8%
Correct	69%
	<hr/> 100%

The authors interpret all answers other than "without merit" as endorsements of the verdict. The rationale for this procedure is not immediately apparent. In cases of disagreement, the judge is given a choice of three characterizations for the jury's action. The first involves narrow ground indeed—the verdict must be without any merit and diametrically at odds with the evidence. The second states that it would be untenable for a judge to reach the conclusion in question, but implies that there may be a reason why a jury might reach such a conclusion. What the "some reason" is, of course, remains unknown. Necessarily it must involve a value judgment, or at least a conclusion based on the judge's experience, about how the jury either should or does operate. A judge may well feel that a verdict is tenable for a jury because he believes juries are generally not very astute and thus more cannot be expected of them. Such an answer need not be, and probably is not, an endorsement of the verdict.

The judge would, in this view of the 14% figure, be critical of the jury's verdict in 23% of all cases—and in 78% of all disagreements. The point, of course, is that either assumption might be justified and, seen in this light, we hardly have any "decisive" word from the judges on how they feel about judge-jury disagreement.

V.

Unfortunately we know little more today about the central issues involved in the traditional jury debate than we knew before *The American Jury*. Though the authors assert that they have explained "in

the large, but with precision" when and why judge and jury disagree, in fact they have told us little more than that judge and jury decide once in four cases differently. As they acknowledged, this datum is not particularly useful. We don't know for certain who is correct—judge or jury. Even if we did, we lack a context in which to decide how much disagreement is too much. Furthermore, we have not been given a convincing explanation of why judges and juries disagree.

The misfortune is that, properly used, the authors' data might have been helpful indeed. Had they used only the answer to Question 34, we would at least have known the judges' explanations for their disagreements with the jury. But the answers to Question 34 were supplemented by a different type of analysis—inferences based on answers to descriptive questions were used in lieu of explanations. The essential problem with the study, then, is that the authors have commingled two very different methods. The judges' direct explanations are adulterated by inferences and explanations provided by the authors. And the authors' inferences based on descriptive questions are useless, because the frequency of occurrence of these characteristics in agreement and disagreement cases is not compared.

It is submitted therefore that *The American Jury*, the most ambitious attempt to date to answer one of the most fundamental, persistent questions in the administration of criminal justice, is a serious failure. Though an effort to replace speculation and conjecture with analysis and fact, *The American Jury* adds little that is new to the jury debate.

A Constitution for Every Man

William Van Alstyne†

Everyman's Constitution. By Howard Jay Graham. Madison: The State Historical Society of Wisconsin, 1968. Pp. xiv, 631. \$12.95.

It may be exceptional to introduce a review of one man's book by beginning with a reference to a different author, but the unorthodoxy will save a great deal of time. Leonard Levy, professor of constitutional history at Brandeis University, has published several justly famous recent works including *The Legacy of Suppression* (an historical treatment of unfree speech), *Jefferson on Civil Liberties*, and *The Origin*

† Professor of Law, Duke University.

of the *Fifth Amendment*. The last, an unhurried review of the evolution of the privilege against self-incrimination, earned a Pulitzer prize last year. Levy's brand is skeptical and liberal, and just now he dominates the field of constitutional historians.

All the more reason, therefore, to pay attention to the Foreword to Howard Jay Graham's *Everyman's Constitution*, as it is written by Professor Levy and it assesses Graham's collected works in the following way:

The year 1968 marks the one hundredth anniversary of the ratification of the Fourteenth Amendment. A most fitting commemoration of that centennial is this collection of essays by Howard Jay Graham, who is surely the greatest authority on the history of the amendment. He is its Maitland, and perhaps our foremost living historian of American constitutional law as well.¹

Howard Jay Graham has played an important part in the developing history of the Fourteenth Amendment. Even as he chronicled its origins and purposes, he influenced its interpretation. By no coincidence, substantive due process of law as the mainstay of decisions against the constitutionality of government regulation came to an end when Graham provided the scholarly proof that the amendment was not designed to benefit business enterprise. Similarly, when he showed that the amendment emerged from the efforts of its framers to ensure that Negroes should have the same rights as other citizens, he provided the historical basis for decisions, which rapidly followed, in support of equal rights regardless of race.²

And that, in brief, is a pretty fair review of *Everyman's Constitution*. The book assembles eleven articles previously published in various law reviews over a period of thirty years, adds two new chapters, and bridges the separations with editorial comment. The result is a smooth and comprehensive treatment of important fourteenth amendment history. Its concern is seemingly with two discrete subjects, but in fact it presents a conscientious historian's brief for one principal theme.

The early chapters revisit the conspiracy theory of the fourteenth amendment that corporations were intended to be protected by the due process clause—a theory most familiarly associated with Roscoe Conkling's argument in *San Mateo County v. Southern Pacific R.R.*,³ and enlarged upon by the books of Charles and Mary Beard who ad-

1. Levy, Foreword to H. GRAHAM, *EVERYMAN'S CONSTITUTION* at vii (1968).

2. *Id.* at vii-viii.

3. 116 U.S. 138 (1885). Conkling's argument is reproduced in full in *Appendix I* to H. GRAHAM, *supra* note 1, at 594-610.

vanced an even broader economic interpretation of the Constitution.⁴ The point of these early chapters, however, is not to analyze the narrow and currently uninteresting problem of whether corporations were in fact secretly intended to be included as persons entitled to substantive due process protection under the fourteenth amendment. It is, rather, to move from a careful examination of the evidence respecting that alleged secret understanding to the broader question of conspiracy: was the fourteenth amendment principally the product of selfish business interests? Was it designed to smuggle in a constitutional basis for protecting economic interests from social regulation by means of judicial review? Were abolitionist concerns, in fact, a mere front that provided the facade but not the real function of the fourteenth amendment?

Graham's powerful (and successful) effort to exorcise this demonic theory is reinforced and complemented by the succeeding chapters on his second subject—the impact of the abolitionists on the formation of the fourteenth amendment. Together with ten Broek's *The Antislavery Origins of the Fourteenth Amendment*,⁵ Graham's one hundred page chapter on the antislavery backgrounds of the fourteenth amendment (which originally appeared in 1950,⁶ a year earlier than ten Broek's work) effectively rehabilitates the humanity of the fourteenth amendment—the amendment was indeed a constitutional commitment to equal rights, precisely as the Supreme Court initially interpreted it before veering away in favor of business interests for the next half century.⁷

Beyond this, Graham's chapter titled *Our "Declaratory" Fourteenth Amendment* does much to aid our understanding of a continuing problem in the use of historical materials to interpret the amendment. Beginning with *Brown v. Board of Education*, the Supreme Court has appeared to encourage the use of historical research in aid of determining whether each alleged form of discrimination was or was not understood to be forbidden by the equal protection clause. In calling for reargument in 1953,⁸ the Court asked counsel to determine whether Congress and the state legislatures contemplated that the fourteenth amendment would abolish segregation in public schools. It also asked

4. See, e.g., 2 C. BEARD & M. BEARD, *THE RISE OF AMERICAN CIVILIZATION* 111-14 (1928).

5. J. TEN BROEK, *THE ANTISLAVERY ORIGINS OF THE FOURTEENTH AMENDMENT* (1951).

6. Graham, *The Early Antislavery Backgrounds of the Fourteenth Amendment* (pts. 1 & 2), 1950 *Wis L. Rev.* 479, 610.

7. See the emphatic dicta of Mr. Justice Strong in *Strauder v. West Virginia*, 100 U.S. 303 (1879) and *Ex Parte Virginia*, 100 U.S. 339 (1879).

8. 345 U.S. 972 (1953).

whether, assuming that the immediate abolition of segregation was not contemplated, the framers nevertheless understood that Congress or the Court would, under future conditions, have power to abolish segregation. Similar references to particular "understandings" were also diligently pursued in the reapportionment cases,⁹ the poll tax¹⁰ and literacy test¹¹ cases, and the antimiscegenation decision.¹² In each instance, the ensuing decision has been subject to considerable criticism—that whatever the liberal virtues of the results, they are insupportable in terms of the framers' original understanding. Mr. Justice Harlan's elaborate dissent in *Reynolds*¹³ is particularly caustic on this point, and the many articles by Alfred Avins¹⁴ carry the criticism forward on every other front.

Graham's writing on "Our 'Declaratory' Fourteenth Amendment," however, effectively supports the wisdom of the second of the two questions posed by the Court in the 1953 *Brown* decision: not the one respecting immediate changes that the framers understood must ensue at once from ratification of the fourteenth amendment, but the one respecting Congress's (and the Court's) prerogative, if any, to enlarge upon those changes under future conditions. In *Brown* itself, the Court concluded conservatively that the evidence on this matter was inconclusive, and so the case was resolved almost entirely on other bases. Graham's writing is more aggressive in this respect. He argues that, in keeping with the open texture of the amendment's language, the animating spirit of the amendment was thematic and ideological, rather than detailed and legislative. The amendment thus embraced a capacity for growth in the particular application of its broad norms, rather than seeking to settle for all time a fixed and legislated answer to each possible controversy in terms of the conditions, perceptions, and information of 1866. To a large extent, this position is shared by Professor Bickel,¹⁵ and by Professor Kelly.¹⁶ Understanding of this view makes less unnerving the statement by Mr. Justice Douglas in *Harper v. Virginia Board of Elections*, that: "Notions of what consti-

9. See, e.g., *Reynolds v. Sims*, 377 U.S. 533 (1964).

10. *Harper v. Virginia Bd. of Elections*, 383 U.S. 663 (1966).

11. *Katzenbach v. Morgan*, 384 U.S. 641 (1966).

12. *Loving v. Virginia*, 388 U.S. 1 (1967).

13. 377 U.S. at 589.

14. See, e.g., Avins, *The Right to Hold Public Office and the Fourteenth and Fifteenth Amendments: The Original Understanding*, 15 U. OF KAN. L. REV. 287 (1967).

15. See Bickel, *The Original Understanding and the Segregation Decision*, 69 HARV. L. REV. 1, 62 (1955).

16. See Kelly, *Clio and the Court: An Illicit Love Affair*, 1965 SUPREME COURT REVIEW 119, 134 (1965); Kelly, *The Fourteenth Amendment Reconsidered*, 54 MICH. L. REV. 1049, 1083-86 (1956).

tutes equal treatment for purposes of the Equal Protection Clause *do* change."¹⁷

Graham does not bemoan the activism of the Supreme Court in its aggressive use of the fourteenth amendment to protect economic concerns. Rather, his is a more measured observation of heavy irony—*viz.*, that the Court was most creative during the first half-century of the fourteenth amendment in areas of least importance to the amendment's origin and animating spirit, even while it was willfully uncreative in areas of the amendment's basic drive. Thus, his suggestion in the preface:

My thesis is simply that what the United States, under these guarantees, did for itself, and for corporations, in curbing manifest and latent hostility and antagonism to corporate enterprise, 1880-1940, the United States can and must do for itself, and for still disadvantaged minorities, using the same techniques and weapons, supplying similar, and, in this case, *intended* process and protection.¹⁸

The thought is not ill-considered as an historian's suggestion in righting history. Much of the Court's early development of substantive due process against local regulation of corporate interests was, after all, the product of Mr. Justice Field whose lack of interest in the amendment's abolitionist underpinnings may be partly understood from a segment of a letter he wrote to Professor Pomeroy, in 1882: "You know I belong to the class, who repudiate the doctrine that this country was made for the people of *all* races. On the contrary, I think it is for our race—the Caucasian race."¹⁹ Yet, far more of the actual formation of the fourteenth amendment was the product of men like Thaddeus Stephens who, dying just three weeks after the fourteenth amendment was proclaimed as ratified, was buried in a plain graveyard in Pennsylvania beneath this epitaph:

I repose in this quiet and secluded spot,
Not from any natural preference for solitude
But, finding other Cemeteries limited as to Race by Charter Rules,
I have chosen this that I might illustrate in my death
The principles which I advocated Through a long life:
EQUALITY OF MAN BEFORE HIS CREATOR.²⁰

Mr. Justice Field having exercised a constitutional compassion for interests of enterprise during the longest tenure of any man in the his-

17. 383 U.S. at 669.

18. H. GRAHAM, *supra* note 1, at ix.

19. H. GRAHAM, *supra* note 1, at 195.

20. F. BRODIE, THADDEUS STEVENS 366 (1959).

tory of the Court, it may not be amiss in the last half of this different century that the Court has also exercised a constitutional compassion of the sort reflected in Thaddeus Stephens' epitaph.

And this, of course, is the explanation of the title of Howard Jay Graham's book. It is not *Everyman's Constitution* in the sense of being a layman's guide through highlights of the Constitution. It is, rather, a more technical but highly readable review of the fourteenth amendment's origins in support of the thesis that the amendment is itself a Constitution meant for *everyone*.

